From the Sensor to Color Images

Olivier Losson, Eric Dinet

Chapter 6

From the Sensor to Color Images

Olivier Losson[a], Éric Dinet[b]

[a] *Laboratoire LAGIS UMR CNRS 8146 – Bâtiment P2*
Université Lille1 – Sciences et Technologies, Villeneuve d'Ascq, France

[b] *Laboratoire Hubert Curien UMR CNRS 5516*
Université Jean Monnet,Saint-Étienne, France

6.1 introduction

In Chapter 5, we saw how color sensors convert a light radiation into an electrical signal to build a digital image. It describes the two currently most widespread sensor technologies (CCD and CMOS) as well as the emerging technologies offered by some manufacturers. These color sensors are nowadays ubiquitous in the objects of everyday life, but the acquisition devices equipped with three sensors are overwhelmingly confined to the professional sectors and to very specific applications. Owing to the complexity of their manufacture, these devices are costly and largely unaffordable by the public. A whole range of so-called "hi-tech" products such as digital cameras, mobile phones, and computers are thus equipped with a single sensor to form a color image, and so are many professional devices such as quality control cameras and video surveillance cameras. One reason for this is the dramatic advances in the operation aiming to obtain a color image from the sensor data. The operation in question, known as *demosaicing*, is the subject of this chapter.

Devices equipped with a single sensor form a color image by estimating it from the so-called *raw* image or color filter array (CFA) image, in which each pixel is only characterized by a single bit of color information. Specifically, the filter mosaic of the CFA samples a single color component (red, green, or blue) at each photoreceptor, and demosaicing aims to estimate the two missing components at the corresponding pixel. This is a far from trivial operation, and the colors estimated thus are less consistent with the color stimuli of the observed scene than those provided by a three-sensor camera. Improving the fidelity of the color image is still a current issue, on which scientists and engineers are working. To obtain an image rendering the colors of the scene as accurately as possible, other processings are typically integrated into the acquisition system, foremost among which are calibration and color correction.

In the following pages, we look at the formation of color images from the data delivered by the sensor, and at the fidelity of these images to the observed scene. The first three sections, largely inspired by the work of Yanqin Yang [YAN 09], deal with the problem of demosaicing, while the last part tackles the problem of color camera calibration. After setting a few notations, in the first section, we present two principles used by the majority of demosaicing methods. The second section presents the key ideas of the main demosaicing algorithms. The issue of fidelity of the estimated image is discussed in the third section, which presents both the main measurement criteria and some results allowing us to select a demosaicing method. Finally, we examine the need for and the implementation of the processing known as *white balance*, usually done before the demosaicing, to obtain a color image that is faithful to the scene regardless of illumination conditions.

6.2. Presentation and formalization of demosaicing

After identifying the need for the demosaicing operation within single-sensor cameras, a formalization is proposed to introduce the notations useful in the following. Then, we introduce the problems associated with color estimation, from a study based on a very simple method using the interpolation of the available levels. This allows us to

set the fundamental principles generally used in demosaicing methods, which is discussed in the next section.

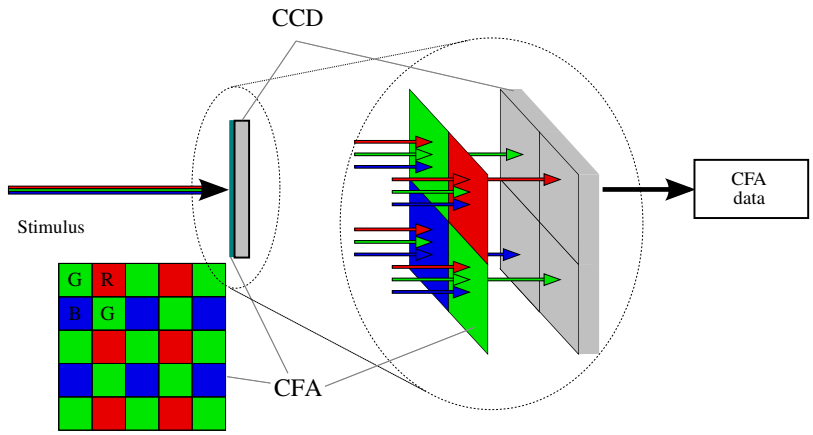
6.2.1. *Need for demosaicing*

As it was seen in Chapter 5, single-sensor color cameras use a mosaic of spectrally selective color filters to sample a single color component at each pixel (see Figure 6.1a). The camera manufacturers have developed several types of filters, but the most widely used remains that of Bayer [BAY 76]. This filter is the subject of most studies and is considered in the following. It features twice as many green filters (G) as red (R) or blue (B) filters, to properly estimate the luminance to which the green component is often equated[1]. The data delivered by the sensor is preprocessed to obtain the CFA image, which therefore contains twice as many pixels whose levels represent the green as pixels representing the red or the blue component. The *demosaicing* operation must then estimate the levels of the two missing components at each pixel in the CFA image to obtain the final color image.

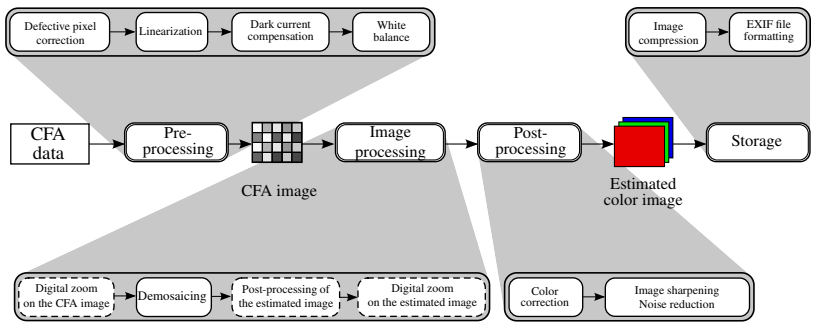
As shown in Figure 6.1b, other operations are typically performed after demosaicing within a single-sensor color camera [LUK 08]. They aim, for instance, to correct the colors, to increase the image sharpness, and to reduce the noise, to render a visually satisfactory color image to the user. All these processes contribute to the quality of the rendered image, and, ultimately, they are the important difference between the different models of digital cameras, as manufacturers and sensor models are few in number. Although the underlying algorithms have some commonalities, the adjustment of their parameters determines the more or less significant presence of residual errors (or *artifacts*) that characterize each camera model.

---

1 The assimilation of the luminance to the component G, used especially to design the Bayer CFA and in the first demosaicing methods, is based on the observation that the curve of the luminous efficiency function of the human visual system in daylight (or *photopic*) vision is similar to that of the colorimetric function of the green primary proposed by the CIE for the reference observer.



(a) Single-CCD technology (simplified), with filters arrangement according to the Bayer CFA



(b) Image acquisition in a single-CCD color camera (detailed diagram). Dotted steps are optional

Figure 6.1. *Structure of a single-CCD color camera*

6.2.2. *Formalization*

A digital image is represented by an array of pixels that we individually note $P(x, y)$, where x and y are the coordinates of the pixel P in the image of size $X \times Y$, with $(x, y) \in \mathbb{N}^2$, $0 \leq x \leq X - 1$ and $0 \leq y \leq Y - 1$. When the image has only one plane, it is written I, and $I(P)$ is the level of each of its pixels. In the case of a color image, the image denoted $\mathbf{I}$ consists of three component planes and $\mathbf{I}(P)$ is a color point in a 3D color space. In the case of the *RGB* space that interests

us, the three components of this vector are $\mathbf{I}(P) = \left(I^R(P), I^G(P), I^B(P)\right)$, where $I^k(P)$, $k \in \{R, G, B\}$, is the level of the color component k at the pixel P. To simplify the notations used for demosaicing and to clarify the spatial coordinates, we similarly adopt in the remainder one of the following simplified notations: $\mathbf{I}(P) = (I^R_{x,y}, I^G_{x,y}, I^B_{x,y}) = (R_{x,y}, G_{x,y}, B_{x,y})$.

To establish a formalism for demosaicing, let us consider the acquisition process of a color image as whether the camera features three sensors or only one sensor (see Figure 6.2). A three-sensor camera (Figure 6.2a) combines the data delivered by the three sensors to form the color image of a scene. This color image $\mathbf{I}$ consists of three component planes I^k, $k \in \{R, G, B\}$. In a single-sensor camera (Figure 6.2b), the formation of a color image is quite different. The CFA image delivered by the sensor has only one plane of levels, and can therefore be represented by a 2D array of integer values (usually between 0 and 255). This image I^{CFA} is demosaiced (taking into account the known configuration of the CFA) to estimate the three component planes of the color image $\hat{\mathbf{I}}$.

Considering the Bayer CFA whose arrangement is shown in Figure 6.3a, the CFA image is defined at each pixel of coordinates (x, y) by the single color component associated with it:

$$I^{CFA}_{x,y} = \begin{cases} R_{x,y} & \text{if } x \text{ is odd and } y \text{ even} \\ G_{x,y} & \text{if } x \text{ and } y \text{ are of the same parity} \\ B_{x,y} & \text{if } x \text{ is even and } y \text{ odd} \end{cases} \quad [6.1]$$

Let $\mathcal{I}^k$, $k \in \{R, G, B\}$, be the subset of pixels of I^{CFA} at which the available component is k (see Figures 6.3b to d):

$$\mathcal{I}^R = \{P(x, y) \in I^{CFA} \mid x \text{ odd and } y \text{ even}\} \quad [6.2]$$

$$\mathcal{I}^G = \{P(x, y) \in I^{CFA} \mid x \text{ and } y \text{ of the same parity}\} \quad [6.3]$$

$$\mathcal{I}^B = \{P(x, y) \in I^{CFA} \mid x \text{ even and } y \text{ odd}\} \quad [6.4]$$

so that $I^{CFA} \equiv \mathcal{I}^R \cup \mathcal{I}^G \cup \mathcal{I}^B$. To determine the color at each pixel $P(x, y)$ of the estimated color image, the demosaicing process (denoted $\mathcal{D}$) generally keeps the color component available at the same coordinates in I^{CFA} and estimates the other two components:

$$I^{CFA}(P) \overset{\mathcal{D}}{\rightarrow} \hat{\mathbf{I}}(P) = \begin{cases} \left(I^{CFA}(P), \hat{I}^G(P), \hat{I}^B(P)\right) & \text{if } P \in \mathcal{I}^R \\ \left(\hat{I}^R(P), I^{CFA}(P), \hat{I}^B(P)\right) & \text{if } P \in \mathcal{I}^G \\ \left(\hat{I}^R(P), \hat{I}^G(P), I^{CFA}(P)\right) & \text{if } P \in \mathcal{I}^B \end{cases} \quad [6.5]$$

In this equation, each triplet of color components contains the one available at the pixel P in I^{CFA} (*i.e.*, $I^{CFA}(P)$) and the other two components estimated by demosaicing (two among $\hat{I}^R(P)$, $\hat{I}^G(P)$, and $\hat{I}^B(P)$).

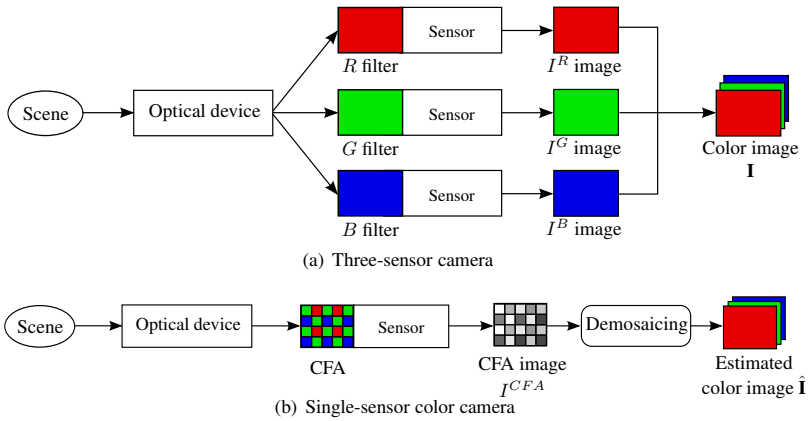
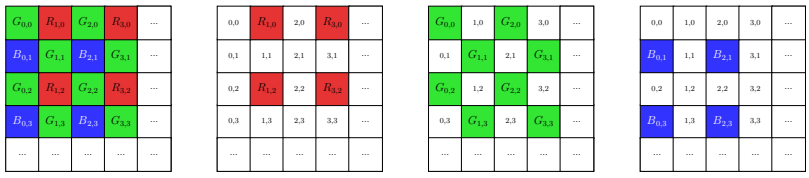


Figure 6.2. *Acquisition of a color image according to the type of camera (simplified diagrams)*

6.2.3. *Implemented principles*

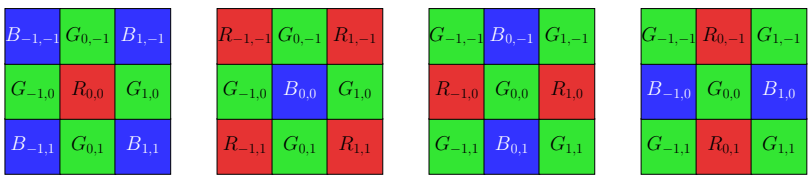
In the CFA image (see Figure 6.3a), there are four different spatial neighborhood structures, shown in Figure 6.4 for a size of 3 × 3 pixels. These are denoted here by the color components available on the center line containing the central pixel of analysis, namely {*GRG*}, {*GBG*}, {*RGR*}, and {*BGB*}. Demosaicing aims to estimate the two

color components missing at the central pixel of each of these four structures, by taking into account the levels of neighboring pixels and the components that are available there. Note that $\{GRG\}$ and $\{GBG\}$ are structurally similar, with the exception that the components R and B are swapped. They can thus be treated in the same way, which also applies to the $\{RGR\}$ and $\{BGB\}$ structures. To denote the levels in these neighborhood structures, we use the notation in relative coordinates of Figure 6.4. Thus, $R_{\delta x,\delta y}$ denotes the level (a red level, in this case) of the pixel at coordinates $(\delta x, \delta y)$ with respect to the central pixel. The coordinates of the central pixel $(0,0)$ are omitted to alleviate notations in the following.



(a) I^{CFA} (b) $\mathcal{I}^R$ (c) $\mathcal{I}^G$ (d) $\mathcal{I}^B$

Figure 6.3. *CFA image and pixel subsets $\mathcal{I}^k$, $k \in \{R,G,B\}$. To clarify the picture, the pixels are colorized with the color component associated with them*



(a) $\{GRG\}$ (b) $\{GBG\}$ (c) $\{RGR\}$ (d) $\{BGB\}$

Figure 6.4. 3×3 *neighborhood structures of pixels in the CFA image. The spatial coordinates are here relative to those of the central pixel*

To introduce the principles generally used in demosaicing methods, let us examine one of the oldest and simplest methods that uses bilinear interpolation, as well as the artifacts it produces. Implemented as an embedded processing since the mid-1980s [COK 86], demosaicing by bilinear interpolation estimates the missing levels at a pixel by averaging the levels available at the nearest neighboring pixels. This is achieved in

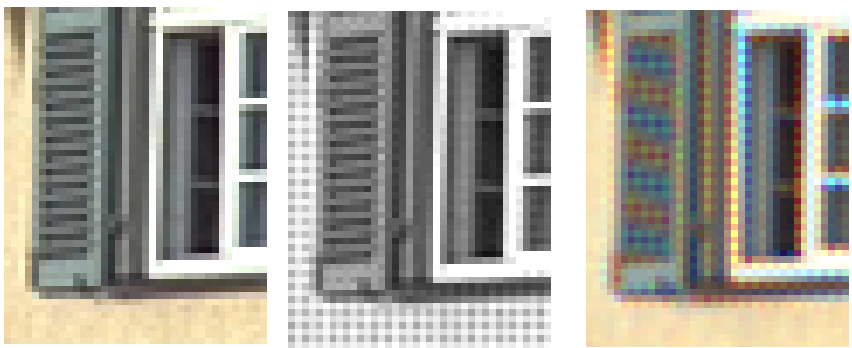
each of the two main directions of the image plane and on each color plane separately. More explicitly, the values of the missing components at the central pixel are estimated using the following equations, according to the structure in question:

$$\text{– for } \{GRG\}: \begin{cases} \hat{B} &= \frac{1}{4}\left(B_{-1,-1} + B_{1,-1} + B_{-1,1} + B_{1,1}\right) \\ \hat{G} &= \frac{1}{4}\left(G_{0,-1} + G_{-1,0} + G_{1,0} + G_{0,1}\right) \end{cases} \quad [6.6]$$

$$\text{– for } \{RGR\}: \begin{cases} \hat{R} &= \frac{1}{2}\left(R_{-1,0} + R_{1,0}\right) \\ \hat{B} &= \frac{1}{2}\left(B_{0,-1} + B_{0,1}\right) \end{cases} \quad [6.7]$$

Figure 6.5 is an example of results obtained by bilinear interpolation. To assess the demosaicing quality, a reference color image is used (typically, an image that comes from a three-sensor camera such as the one in Figure 6.5a). Its color components are sampled according to the CFA mosaic to form the CFA image (Figure 6.5b). This one is then demosaiced, and the obtained estimated image (Figure 6.5c) can then be compared to the reference image. Demosaicing by bilinear interpolation is simple and fast, but although it provides satisfactory results in image regions of homogeneous colors, it generates erroneous colors at many pixels in regions with high spatial frequencies.

To precisely investigate the causes of these artifacts, let us simulate their generation using a synthetic image, like Chang and Tan [CHA 06].

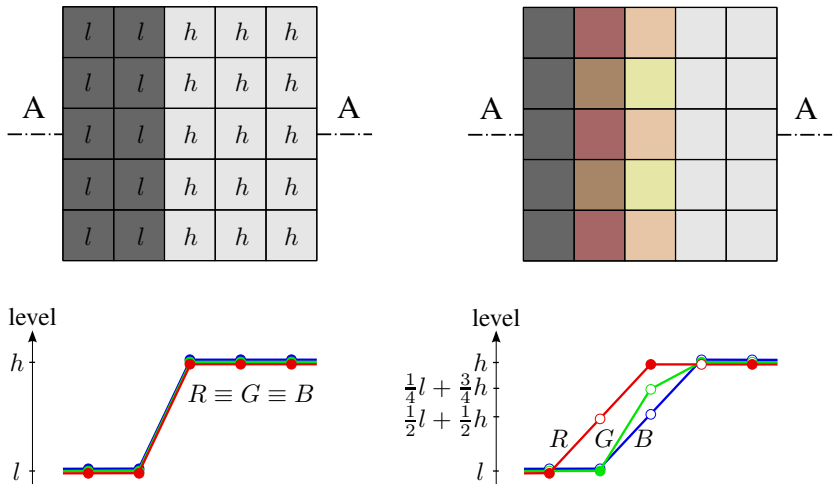


(a) Reference image $\mathbf{I}$ (b) CFA image I^{CFA} (c) Estimated image $\hat{\mathbf{I}}$

Figure 6.5. *Example of demosaicing by bilinear interpolation on a sample taken from the "Houses" image from the Kodak database [KOD 91]*

In the reference image shown in Figure 6.6a, two homogeneous regions are separated by a vertical transition, which reproduces the boundary between two real objects characterized by different shades of gray. The three color components of each pixel are thus equal. The levels of pixels (labeled as l) representing the dark object on the left are lower than those of pixels (labeled as h) representing the bright object on the right. Demosaicing the corresponding CFA image by bilinear interpolation yields the result illustrated in Figure 6.6b. On the color planes R and B, this algorithm produces a column of pixels of intermediate levels, whose values are the averages of the levels representing the two objects. On the green plane, however, it generates an alternating pattern of pixels with two intermediate levels around the boundary, one of low value $\frac{3}{4}l + \frac{1}{4}h$ and the other of high value $\frac{1}{4}l + \frac{3}{4}h$. Comparison of the marginal profiles of the center line of pixels of both images reveals that the transition located at the same horizontal positions for the three components of the reference image is not reproduced identically for the three components of the estimated image. This inconsistency generates false colors in the estimated image. On each color plane of the reference image, the transition corresponds to a sudden break of homogeneity along its normal direction. After bilinear interpolation, averaging the component levels of pixels located on both sides of the transition has the effect of making it less sharp. Although established on an achromatic image, these findings help to highlight two fundamental properties of color images that must be respected to improve the result of demosaicing: *spectral correlation* and *spatial correlation*.

The property of *spectral correlation* has been studied by Gunturk *et al.* [GUN 02]. These authors show that the levels of the three components are highly correlated in a natural color image, and especially in areas of high spatial frequencies. To exploit this spectral intra-pixel correlation for demosaicing, two principles are mainly used in the literature: the local constancy of the *ratio* of color components, and the local constancy of their *difference*. Historically, the first method implementing spectral correlation is that of Cok [COK 87], which uses the principle of local constancy of the "hue" (understood as the ratio of the chrominance and luminance and equated to R/G or B/G). After estimating all the missing green levels by bilinear interpolation, this



(a) Reference image **I** (b) Estimated image $\hat{\mathbf{I}}$

Figure 6.6. *Marginal images and profiles of levels of pixels located on the center line A-A, for a reference image (a) and the corresponding image estimated by bilinear interpolation (b). In the profiles, the color dots denote available levels and the white dots denote estimated levels*

method estimates the missing red (respectively blue) levels by weighting the green level with the average "hue" values of neighboring pixels where the red (respectively, blue) component is available. For instance, to interpolate the blue level at the center of the {*GRG*} structure (see Figure 6.4a), the following equation is applied. It uses the four diagonal neighbors at which the blue component is available:

$$\hat{B} = \hat{G} \cdot \frac{1}{4}\left[\frac{B_{-1,-1}}{\hat{G}_{-1,-1}} + \frac{B_{1,-1}}{\hat{G}_{1,-1}} + \frac{B_{-1,1}}{\hat{G}_{-1,1}} + \frac{B_{1,1}}{\hat{G}_{1,1}}\right] \qquad [6.8]$$

Such a bilinear interpolation of the color component ratios is based on the fact that this ratio is locally constant in a homogeneous region, which is justified by Kimmel [KIM 99] under the assumption of Lambertian observed surfaces. Another simplified model of the inter-channel correlation, also widely used in the literature, is based on the local constancy principle of the *difference* of color components (see justification in [LIA 07]). The interpolation step of the chrominance by

Cok's method is then rewritten using the component difference average, and equation [6.8] becomes:

$$\hat{B} = \hat{G} + \frac{1}{4}\Big[(B_{-1,-1} - \hat{G}_{-1,-1}) + (B_{1,-1} - \hat{G}_{1,-1}) + (B_{-1,1} - \hat{G}_{-1,1}) + (B_{1,1} - \hat{G}_{1,1})\Big] \quad [6.9]$$

The two local consistency principles of the ratio and of the difference of color components are generally consistent, but that using the ratio is more prone to generate artifacts, particularly in the image regions where the components R and/or B are saturated.

The *spatial correlation* property is easily understood by considering a color image as composed of juxtaposed homogeneous regions. Within a homogeneous region, all pixels are characterized by similar levels, and this, for each color component. To estimate the missing levels of each pixel in question, it is thus possible to use the levels of neighboring pixels. However, that is more difficult at pixels located near transitions between two distinct regions, where the local variation of color components is strong. From the demosaicing point of view, the respect of the spatial correlation property thus incites to avoid interpolating the missing components of a pixel using neighboring pixels that do not belong to the same homogeneous region. This principle is implemented in many demosaicing methods, and the next section presents a few examples. It is typically used–sometimes combined with that of spectral correlation–in a first step of estimating the green component, since the green information is the most dense in the CFA image and represents the luminance of the image to be estimated (following Bayer's idea). The estimation of the red and blue components (considered the chrominance according to Bayer) is done only in a second step, using the already interpolated luminance and the spectral correlation property.

6.3. Demosaicing methods

The problem of estimating the levels missing in the CFA image arose in the late 1970s, right after the invention of the first filter mosaics for

single-sensor color cameras. Numerous demosaicing algorithms have been developed since then, and it is impossible to draw up an exhaustive list here. Highly comprehensive inventories are available in the literature (such as [LI 08, LUK 08, LOS 10]) and online[2]. After the explanation of the basics in the previous section, we now introduce some of the most significant methods using a spatial analysis. The second subsection focuses on the principle of frequency analysis, used by the currently most efficient algorithms. We briefly present other approaches proposed for demosaicing, as well as post-processing typically implemented to improve the quality of the estimated image.

6.3.1. *Methods based on a spatial analysis*

Noting that the method based on the hue constancy suffers from serious estimation errors in regions of high spatial frequencies (see Figure 6.9b), Cok proposed the first edge-sensitive demosaicing algorithm [COK 94]. Based on pattern recognition, this method improves the green-level estimation by classifying (as *edge*, *stripe* or *corner*) the 3×3 neighborhood of pixels where that level is missing, then by adapting the interpolation formula according to the neighborhood class. The criterion used to distinguish the three classes, which compares the neighboring levels to their average, is very simple and not always sufficient to correctly identify the actual shape. However, the idea can be developed further, and it [CHA 06] marks a milestone because it opens the way to methods adapting themselves to the local properties of the image.

Many methods thus implement the property of spatial correlation in demosaicing. Most of them exploit the principle of interpolating only levels whose pixels belong to the same homogeneous region. In other words, when the neighborhood of a pixel in question is located on the transition between two homogeneous regions, it is necessary to estimate the values missing at this pixel *along* the transition and not through it. A key point is thus to correctly determine the transition direction from the samples available in the CFA image. Computing a gradient

2 http://www.danielemenon.netsons.org/top/demosaicking-list.php.

is a straightforward solution to this problem. The method proposed by Hibbard [HIB 95] uses horizontal and vertical gradients to determine the direction along which the interpolation provides the best estimation of the green level. As an example, let us consider the {*GRG*} structure again (see Figure 6.4a), in which the green level $\hat{G}$ of the central pixel is estimated in two steps. The first is to compute an approximation of the gradient norm (hereafter referred to as *gradient* for simplification purposes) according to the horizontal and vertical directions:

$$\Delta^x = \left|G_{-1,0} - G_{1,0}\right| \quad [6.10]$$

$$\Delta^y = \left|G_{0,-1} - G_{0,1}\right| \quad [6.11]$$

The second step is to interpolate the green level using the equation:

$$\hat{G} = \begin{cases} (G_{-1,0} + G_{1,0})/2 & \text{if } \Delta^x < \Delta^y \quad [6.12a] \\ (G_{0,-1} + G_{0,1})/2 & \text{if } \Delta^x > \Delta^y \quad [6.12b] \\ (G_{0,-1} + G_{-1,0} + G_{1,0} + G_{0,1})/4 & \text{if } \Delta^x = \Delta^y \quad [6.12c] \end{cases}$$

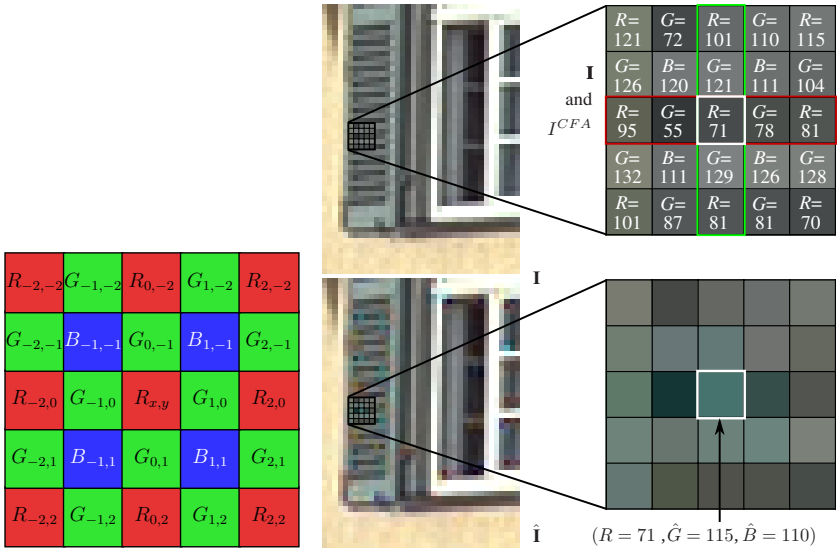
Laroche and Prescott [LAR 93] propose a variant of this method that takes into account the levels available in a 5×5 neighborhood to determine the transition direction, for example $\Delta^x = \left|2R - R_{-2,0} - R_{2,0}\right|$. Hamilton and Adams [HAM 97] combine the two previous approaches, using a first-order differentiation for the green component and a second-order differentiation for the red and blue components. For instance, to estimate the green level in the case of the {*GRG*} structure (see Figure 6.7a), this method first computes the following horizontal and vertical differences:

$$\Delta^x = \left|G_{-1,0} - G_{1,0}\right| + \left|2R - R_{-2,0} - R_{2,0}\right| \quad [6.13]$$

$$\Delta^y = \left|G_{0,-1} - G_{0,1}\right| + \left|2R - R_{0,-2} - R_{0,2}\right| \quad [6.14]$$

then interpolates the green level using the equation:

$$\hat{G} = \begin{cases} \left(G_{-1,0} + G_{1,0}\right)/2 + \left(2R - R_{-2,0} - R_{2,0}\right)/4 & \text{if } \Delta^x < \Delta^y \quad [6.15a] \\ \left(G_{0,-1} + G_{0,1}\right)/2 + \left(2R - R_{0,-2} - R_{0,2}\right)/4 & \text{if } \Delta^x > \Delta^y \quad [6.15b] \\ \left(G_{0,-1} + G_{-1,0} + G_{1,0} + G_{0,1}\right)/4 & \\ \quad + \left(4R - R_{0,-2} - R_{-2,0} - R_{2,0} - R_{0,2}\right)/8 & \text{if } \Delta^x = \Delta^y \quad [6.15c] \end{cases}$$



(a) 5×5 CFA neighborhood centered in R

(b) Example of demosaicing result

Figure 6.7. *Hamilton and Adams' method [HAM 97]. (a) One of the CFA neighborhood structures used by this method (size* 5×5 *pixels, centered at R). (b) Exampleof demosaicing result, on the same image sample as in Figure 6.5. The* 5×5 *detail has the CFA neighborhood structure of figure (a) and the corresponding levels in the CFA image are represented at the upper right, superimposed on the reference image. This detail highlights the failure of the method to correctly determine the interpolation direction: at its central pixel, since* $\Delta^x = 57 > \Delta^y = 48$*, the vertical neighboring pixels are erroneously used to estimate* $\hat{G}$ *according to equation* [6.15b] *(see detail of the estimated image at the lower right)*

This algorithm yields much better results than Hibbard's method, not only because it computes the gradient more accurately by combining the information from two color components, but also because it exploits the spectral correlation to interpolate the green component[3]. However, the criterion used to determine the interpolation direction may be inappropriate and may provide unsatisfactory results in regions

3 The horizontal estimation of the green level (equation [6.15a]) can be rewritten as the average of an interpolation term on the left $\hat{G}^l = G_{-1,0} + \frac{1}{2}(R - R_{-2,0})$ and of an interpolation term on the right $\hat{G}^r = G_{1,0} + \frac{1}{2}(R - R_{2,0})$. Each of these terms expresses the local constancy of the component difference, since and $\hat{G}^l - R = G_{-1,0} - \frac{1}{2}(R_{-2,0} + R)$ and $\hat{G}^r - R = G_{1,0} - \frac{1}{2}(R + R_{2,0})$.

containing thin objects or textures. For instance, Figure 6.7b shows that the approximation of horizontal and vertical differences (Δ^x and Δ^y) does not always lead to the right decision for the interpolation direction.

Wu and Zhang [WU 04] propose an approach to determine this direction more reliably, again using a local neighborhood. Two candidate levels are computed to interpolate the missing green value at a pixel: one is determined with the horizontal neighbors, and the other with the vertical neighbors. Then, the missing value R or B is estimated according to the horizontal and vertical directions with each of these two candidates for the G. Finally, the selected interpolation direction is that for which the component differences $(R-G)$ and $(B-G)$ show minimal variations. This interpolation direction allows us to select the levels (previously computed) to take into account when estimating the missing components of the pixel, thus ensuring the consistency of interpolation directions between these components. Wu and Zhang's method uses the same formula as that of Hamilton and Adams to interpolate the missing color levels, but improves the determination of the interpolation direction by using a 3×3 neighborhood rather than a single row or column, and by measuring the gradient of the component *differences* over this neighborhood.

Other authors attempt to refine the selection of the interpolation direction to still more accurately estimate the pixels representing the observed scene. For example, Hirakawa and Parks [HIR 05] propose a selection criterion using the number of pixels with homogeneous colors in the neighborhood of a given pixel. By computing the color distances between the pixel and its neighbors in the CIE space $L^*a^*b^*$, better suited for the representation of human perception of colors than the *RGB* space, these authors propose an adaptive thresholding homogeneity criterion that allows us to reduce the color artifacts due to an improper selection of the interpolation direction. Chung and Chan [CHU 06], after showing that the interpolation of the green plane is crucial to the quality of the estimated image, propose to estimate the variance of the color component difference in a neighborhood to evaluate the local homogeneity, and to choose the direction corresponding to the

minimum variance. This criterion is used to refine the estimation of the green component, especially in textured regions.

The methods mentioned above, whether Cok's based on pattern recognition or those using a directional gradient, require a classification step of the neighborhood to perform the interpolation. Kimmel [KIM 99] proposes a linear interpolation with adaptive weights that merges these two steps. To each available level of the neighborhood, this method associates a normalized weight that depends on a directional gradient. The direction of that gradient is specific to each neighboring pixel. For instance, to interpolate the green level at the center of the $\{GRG\}$ or $\{GBG\}$ structure, the equation is the following:

$$\hat{G} = \frac{w_{0,-1} \cdot G_{0,-1} + w_{-1,0} \cdot G_{-1,0} + w_{1,0} \cdot G_{1,0} + w_{0,1} \cdot G_{0,1}}{w_{0,-1} + w_{-1,0} + w_{1,0} + w_{0,1}} \quad [6.16]$$

where the coefficients $w_{\delta x,\delta y}$ are the weights computed over the neighborhood of the pixel in question. To exploit spatial correlation, these weights are adjusted to reflect the shape found in the neighborhood. Thus, the interpolation automatically adapts to the transition present in the image. Of course, the determination of these weights is crucial to the quality of results provided by this method, and several authors [LU 03, LUK 05] propose improvements to the formula originally used by Kimmel for their computation.

Finally, another interpolation method without neighborhood classification is worth mentioning. This original method is based on the fact that demosaicing by interpolation exhibits strong similarities with the super-resolution problem approach. With Orchard, Li [LI 01] achieves demosaicing by adapting the algorithm he proposes in his PhD thesis [LI 00] to increase the resolution of a gray scale image. In both problems, conventional approaches by interpolation (bilinear and bicubic) smooth the transitions and produce artifacts in regions of high spatial frequencies. Li's method exploits spatial correlation by evaluating a local covariance of the levels to interpolate the missing values with no directional gradient computation. To achieve demosaicing, each subset $\mathcal{I}^k$, $k \in \{R, G, B\}$, from the CFA image is considered as a sub-sampling of the corresponding plane $\hat{I}^k$ of the image to be

estimated. The main problem is to locally compute the covariance of the levels in the high-resolution image from the known levels in the low-resolution image. This is possible by exploiting the *geometric duality* principle: since the covariance is computed over a local neighborhood in the low-resolution image, the equivalent high-resolution covariance is estimated by geometric duality by considering pairs of pixels at the two resolutions according to the same direction of the image plane.

To close this section about methods based on spatial analysis, let us recall that the green plane is often interpolated first. This estimation is essential because the green component generally carries most of the high-frequency information, particularly in edges or texture regions of natural images. Once the green plane is completely determined, it is used to estimate the chrominance. Hence, there is need to select a method that successfully exploits spatial correlation. A detailed study of the methods mentioned above [LOS 10] shows, in addition, that those estimating the missing green levels only from CFA green levels (case of Cok's [COK 86], Li and Orchard's [LI 01], and bilinear interpolation methods) generally yield poorer results than others. The estimation of the green plane is improved when using the information from the R and B components. A powerful demosaicing method should therefore make the most of spatial and spectral correlations, simultaneously and for each color component.

6.3.2. *Methods based on a frequency analysis*

A demosaicing approach using the frequency domain[4] is proposed by Alleysson *et al.* [ALL 05]. This is the origin of a particularly important family of demosaicing algorithms, as they are currently the most efficient ones (see next section). Their principle is to represent a CFA image as a combination of a luminance[5] component at low spatial

4 The frequency mentioned here is the *spatial* frequency (in cycles/pixel) that is defined as the inverse of the number of adjacent pixels representing the same series of levels, along a preferred direction in the image (typically, the horizontal or vertical direction).

5 The "luminance" term denotes the achromatic component. It is re-used here from publications about demosaicing methods by frequency analysis.

frequencies and of two chrominance components modulated at high spatial frequencies[6], and then to estimate the color image by adequately selecting the frequencies. To outline such an approach, we retain the formalism proposed by Dubois [DUB 05].

Let us assume that each component $k \in \{R, G, B\}$ of a color image corresponds to an underlying signal f^k. Demosaicing then involves computing an estimation $\hat{f}^k$ (coinciding with $\hat{I}^k$) at each pixel. Let us also assume that there exists a signal, denoted as f^{CFA} and called *CFA signal*, underlying the CFA image (*i.e.* coinciding with I^{CFA} at each pixel). The CFA signal value at each pixel of coordinates (x, y) can be expressed as the sum of the spatially sampled signals f^k:

$$f^{CFA}(x,y) = \sum_{k=R,G,B} f^k(x,y) m^k(x,y) \qquad [6.17]$$

where $m^k(x, y)$ is the sampling function of the component k corresponding to the Bayer CFA shown in Figure 6.3a:

$$m^R(x,y) = \tfrac{1}{4}\left[1-(-1)^x\right]\left[1+(-1)^y\right] \qquad [6.18]$$

$$m^G(x,y) = \tfrac{1}{2}\left[1+(-1)^{x+y}\right] \qquad [6.19]$$

$$m^B(x,y) = \tfrac{1}{4}\left[1+(-1)^x\right]\left[1-(-1)^y\right] \qquad [6.20]$$

By setting $\begin{bmatrix} f^L \\ f^{C_1} \\ f^{C_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} f^R \\ f^G \\ f^B \end{bmatrix}$, the expression of f^{CFA} becomes:

$$f^{CFA}(x,y) = f^L(x,y) + f^{C_1}(x,y)\underbrace{(-1)^{x+y}}_{e^{j2\pi(x+y)/2}} + f^{C_2}(x,y)\underbrace{\left[(-1)^x - (-1)^y\right]}_{e^{j2\pi x/2} - e^{j2\pi y/2}} \qquad [6.21]$$

6 Bayer's hypothesis is not used in this representation. Luminance is not assimilated to the green component, nor chrominance to the red and blue components.

The CFA signal can then be interpreted as the sum of a luminance component f^L at baseband, of a chrominance component f^{C_1} modulated at spatial frequencies (horizontal and vertical) $(u = 0.5, v = 0.5)$, and of another chrominance component f^{C_2} modulated at spatial frequencies $(u = 0.5, v = 0)$ and $(u = 0, v = 0.5)$[7]. If it is possible to estimate the functions f^L, f^{C_1}, and f^{C_2} from the CFA signal at each pixel, the estimated levels of the color components R, G, and B are then simply found as:

$$\begin{bmatrix} \hat{f}^R \\ \hat{f}^G \\ \hat{f}^B \end{bmatrix} = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \hat{f}^L \\ \hat{f}^{C_1} \\ \hat{f}^{C_2} \end{bmatrix} \qquad [6.22]$$

The Fourier transform of the CFA signal can be expressed from equation [6.21] as:

$$\begin{aligned} F^{CFA}(u\ ,\ v) = & F^L(u\ ,\ v) + F^{C_1}(u - 0.5, v - 0.5) \\ & + F^{C_2}(u - 0.5, v) - F^{C_2}(u, v - 0.5) \end{aligned} \qquad [6.23]$$

where the terms are the Fourier transforms of $f^L(x, y)$, of $f^{C_1}(x, y)(-1)^{x+y}$, as well as of the two signals defined as $f^{C_{2a}}(x, y) = f^{C_2}(x, y)(-1)^x$ and $f^{C_{2b}}(x, y) = -f^{C_2}(x, y)(-1)^y$.

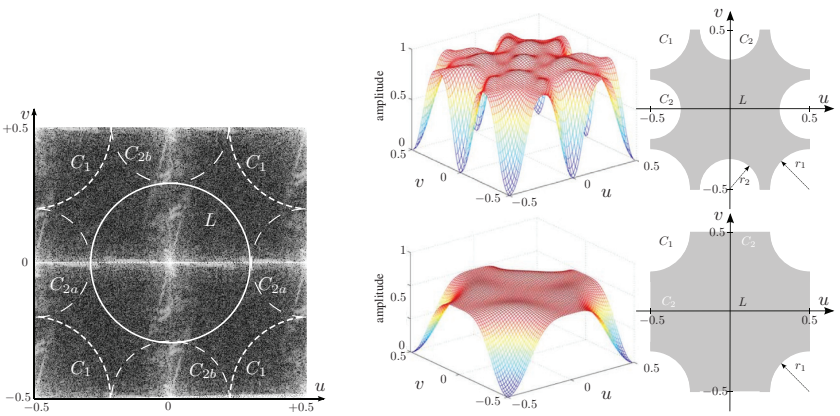
Observing the energy distribution of a CFA image in the frequency plane (see example in Figure 6.8a) reveals a concentration in nine quite distinct regions, centered on the spatial frequencies corresponding to equation [6.23]. In particular, the energy of $F^{C_2}(u - 0.5, v)$ lies on the horizontal frequency axis u and that of $F^{C_2}(u, v - 0.5)$ lies on the vertical frequency axis v. The energy of $F^L(u, v)$ is mainly concentrated at the center of the frequency plane, while that of $F^{C_1}(u - 0.5, v - 0.5)$ is located in the diagonal regions (called "corners") of the plane. The key of methods based on frequency selection thus lies in the design of filters allowing us to effectively separate the components of luminance L and of chrominance C_1 and C_2. The bandwidth of these filters should be

7 It is easy to verify that, on a gray scale image for which $f^R = f^G = f^B$, the two chrominance components of the CFA signal are zero.

adjusted with particular care, given the mutual overlapping (*aliasing*) of the three signal spectra.

Various proposals of selective filters have emerged. For instance, the original article of Alleysson *et al.* [ALL 05] uses the filter whose spectrum is reproduced on the upper part of Figure 6.8b to isolate the luminance component. Dubois [DUB 05] proposes to form the estimation of f^{C_2} by giving more weight to the subcomponent (C_{2a} or C_{2b}) that is less prone to spectral aliasing with the luminance. In a local region of the image, such aliasing occurs mainly in either the horizontal direction or the vertical direction.

Lian *et al.* [LIA 05] show that the estimation of the luminance by these methods is sensitive to the bandwidth of the selective filters. Yet, the parameters of these filters (r_1 and r_2 in Figure 6.8b) depend on the image content and are difficult to adjust. In addition, the luminance selection by low-pass filtering the CFA image causes the loss of the high-frequency information located along the horizontal and vertical directions, to which the human visual system is particularly sensitive.



(a) Energy of a CFA image (b) Filters used to estimate the luminance

Figure 6.8. *Frequency analysis of a CFA image (frequencies in cycles/pixel). (a) Distribution, in the frequency plane, of the normalized energy (Fourier transform module) of the "Lighthouse" CFA image (number 19 of the Kodak database) [ALL 05]. (b) Filters (spectrum and bandwidth) used to estimate the luminance, proposed by Alleysson* et al. *[ALL 05] (top) and Lian* et al. *[LIA 07] (bottom)*

The same authors [LIA 07] point out that along these directions, the spectrum associated with the filter used to estimate the component C_2 is zero if the G locations of the CFA are solely considered. At these locations, they hence apply a filter (see bottom of Figure 6.8b) whose spectrum is zero in the "corners" of the frequency plane where the energy of C_1 is concentrated, which provides the spectrum of the luminance F^L at these G locations. The advantage of doing so is that L exhibits less spectral overlapping with C_1 than with C_2, as it can be seen on the example in Figure 6.8a. To estimate the luminance at R and B locations of the CFA, as it is difficult to isolate the component C_2, the authors use a spatial analysis based on the component difference constancy (exploiting spectral correlation) and a linear interpolation with adaptive weighting (exploiting spatial correlation).

6.3.3. *Other methods and post-processing*

This section is a supplement to the presentation of demosaicing methods. We very briefly talk about the classical post-processing performed by the algorithms already mentioned to remove demosaicing artifacts, as well as alternative approaches.

A post-processing is often performed on the estimated image to correct the estimated colors, usually by increasing iteratively the spectral correlation among the three color components. The median filter, typically used to remove impulse noise in a gray scale image, was historically highly used for its ability to effectively remove false colors without damaging the local variations. Freeman [FRE 88] is the first to take advantage of this filter by applying it to the estimated planes of component differences $R - G$ and $B - G$ (that generally contain few high spatial frequencies), which improves the bilinear interpolation estimation quite significantly. Several authors also include this filter to remove the demosaicing artifacts, such as Hirakawa and Parks [HIR 05], Lu and Tan [LU 03] or Chang and Tan [CHA 06]. A common approach to these last two works is to pre-identify the image regions that may contain artifacts, then to apply the median filter only on those regions. Among other interesting corrective procedures, the reader can refer to that incorporated

into (previously described) Kimmel's method [KIM 99], or to that of Menon *et al.* [MEN 06] taking into account the local edge direction.

Demosaicing methods known as "by regularization" may be viewed as sophisticated correction approaches of the estimated colors. They indeed require an initial estimation $\tilde{\mathbf{I}}$ generated by a simple method (for example, Hamilton and Adams'), that is then iteratively improved by exploiting the spectral correlation principle. In this way, Keren and Osadchy [KER 99] apply a regularization by minimizing (with a finite element method) a cost function that includes a spatial smoothing term and a color correlation term:

$$\begin{aligned}\text{Cost} = &\int\int \sum_{k=R,G,B} \left[\left(\frac{\partial^2 \tilde{I}^k}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 \tilde{I}^k}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 \tilde{I}^k}{\partial y^2} \right)^2 \right] dxdy \\ &+ \gamma \int\int \left(\tilde{\mathbf{I}}_{x,y} - \bar{\mathbf{I}}_{x,y} \right) \mathbf{C}_{x,y}^{-1} \left(\tilde{\mathbf{I}}_{x,y} - \bar{\mathbf{I}}_{x,y} \right)^t dxdy \end{aligned} \quad [6.24]$$

where $\bar{\mathbf{I}}_{x,y}$ and $\mathbf{C}_{x,y}$ are the average color and the color covariance matrix respectively at pixel (x, y), and where γ is a positive constant. Gunturk *et al.* [GUN 02] propose a method that gradually homogenizes the high-frequency characteristics among the three component planes, while keeping the data available in the CFA image. These objectives are fulfilled thanks to two convex sets of constraints on which the algorithm alternately projects the estimated data. The first set, called "Observation", ensures the consistency with the data available in the CFA image. The second, called "Detail", is based on a decomposition of each plane R, G, and B into four frequency sub-bands resulting from the filter bank approach. This algorithm yields excellent demosaicing results, and it has been considered as a performance comparison benchmark for many years. However, it is complex and relatively expensive in computing time (see Table 6.1), and the final quality partly depends on that of the original estimation. A summary of existing regularization approaches is available in the article of Menon *et al.* [MEN 09].

A last important family of methods, following Taubman's idea [TAU 00], considers demosaicing as a generalized inverse

problem [RIB 08]. Here, we take up Chaix de Lavarène's [CHA 07] formulation and notations, regarding images as random variables in this Bayesian framework. The CFA image X is viewed as the transform of the reference color image $\mathbf{Y}$ by the sampling process, represented as a projection operator $\mathbf{Pr}$, that is $X = \mathbf{Pr}\,\mathbf{Y}$. Assuming that the reference data is a linear combination of the acquired CFA data, and that the acquisition process is spatially invariant, the estimation problem can be solved by a Wiener approach. The demosaicing problem thus boils down to finding a matrix $\mathbf{D}$ allowing us to estimate the color image $\hat{\mathbf{Y}} = \mathbf{D}\,X$ from the CFA image X, while minimizing the mean square error to the reference image $e = E[\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2]$. Wiener's solution is written as:

$$\mathbf{D} = \left(E[\mathbf{Y}\,X^t]\right)\left(E[X\,X^t]\right)^{-1} \qquad [6.25]$$

where $E[\cdot]$ represents the mathematical expectation. Note that this equation requires a dataset of reference images $\mathbf{Y}$ that are representative of images that will be demosaiced later using the matrix $\mathbf{D}$. This inverse problem approach also allows us to perform other operations in parallel with the demosaicing. For instance, Condat [CON 10] uses it to perform a demosaicing–denoising by a variational approach that minimizes the total variation under the constraint of consistency with the acquired data. Soulez and Thiébaut [SOU 09] jointly achieve deblurring and demosaicing. To be as complete as possible on the resolution of this underdetermined problem of demosaicing, let us finally mention that some methods use neural networks [KAP 00] or Hidden Markov fields [MUK 01].

6.4. Quality of the estimated image

Evaluating the quality of the estimated image (and thus the performance of the demosaicing method used) can be achieved in various ways: subjectively or objectively, with or without reference, using different metrics, etc. (see Chapter 9 on the evaluation of the quality of color images). We focus here on the objective evaluation of the estimated image by assuming the availability of the reference image, as

it is performed in the vast majority of articles on demosaicing[8]. Several metrics have been proposed for this evaluation according to the objective. In addition to conventional fidelity metrics based on the *RGB* color space, there are:

– metrics based on perceptually uniform spaces, designed to better reflect the quality of an image as the human visual system perceives it [CHU 06, ZHA 97];

– metrics based on the analysis of the artifacts generated in the estimated image, designed to assess their impact on low-level processing applied later on this image [LU 03, YAN 09].

In the context of color image formation, we are solely interested in *fidelity* metrics of the estimated image relative to the reference image. The first subsection presents the usual procedure and criteria used to carry out this evaluation. Results are presented in the second subsection, which allow us to discuss the performance of the different demosaicing methods detailed in this chapter.

6.4.1. *Fidelity criteria of the estimated image*

To objectively assess the quality of the demosaicing result, the experimental protocol already mentioned in the previous section is still followed. The CFA image is first simulated from the reference color image by selecting only one color component among the three according to the CFA mosaic arrangement. A demosaicing method is next applied to the CFA image to provide the estimated image. The objective evaluation of the demosaicing quality is then based on the pixel-to-pixel comparison between the reference image and the estimated image. Classically, objective evaluation metrics sum up the errors between the pixel levels in both images. At each pixel, the error between the reference image and the estimated image is quantified using a color distance between two color points in a 3D space.

8 To our knowledge, only the article of Longère *et al.* [LON 02] describes the experiment of a subjective assessment of the demosaicing quality, involving human observers and regulated by a series of experimental protocols.

Classical fidelity criteria are restricted to the *RGB* space. The demosaicing literature uses three main criteria, the last being deduced from the second:

1) Mean absolute error (*MAE*).

The *MAE* criterion estimates the mean absolute error between the reference image $\mathbf{I}$ and the estimated image $\hat{\mathbf{I}}$ as:

$$MAE(\mathbf{I}, \hat{\mathbf{I}}) = \frac{1}{3XY} \sum_{k=R,G,B} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \left| I_{x,y}^k - \hat{I}_{x,y}^k \right| \quad [6.26]$$

Let us recall that $I_{x,y}^k$ denotes the level of the color component k at the pixel located at spatial coordinates (x, y) in image $\mathbf{I}$, X and Y being the numbers of columns and rows of that image. The value of the *MAE* criterion varies between 0 and 255, and the lower it is, the better is the demosaicing.

2) Mean square error (*MSE*).

The *MSE* criterion estimates the mean square error between the reference image $\mathbf{I}$ and the estimated image $\hat{\mathbf{I}}$ as:

$$MSE(\mathbf{I}, \hat{\mathbf{I}}) = \frac{1}{3XY} \sum_{k=R,G,B} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (I_{x,y}^k - \hat{I}_{x,y}^k)^2 \quad [6.27]$$

The value range for this metrics is $\left[0, 255^2\right]$ and the interpretation of its values is identical to that of the *MAE* criterion: the optimal demosaicing quality corresponds to the zero value of the *MSE* criterion.

3) Peak signal-noise ratio (*PSNR*).

The *PSNR* is a set of distortion metrics particularly used in image compression. It quantifies the encoder performance by measuring the reconstruction quality of the compressed image relative to the original image. Many authors (*e.g.* [ALL 05, HIR 05, LIA 07, WU 04]) apply this criterion to quantify the demosaicing performance. Usually expressed in decibels, its definition is:

$$PSNR(\mathbf{I}, \hat{\mathbf{I}}) = 10 \cdot \log_{10} \left(\frac{d^2}{MSE} \right) \quad [6.28]$$

where d is the maximum possible level for the two compared images. In the standard case of an image whose color components are coded on 8 bits, d equals 255. The better is the demosaicing quality, the higher is the *PSNR* value. The *PSNR* value of images estimated by classical demosaicing methods usually varies between 30 and 40 dB, which corresponds to values between 65.03 and 6.50 for the *MSE* criterion.

Note that these three criteria can also be used to measure the errors on each color component plane. On the green plane, for instance, the mean square error is expressed as:

$$MSE^G(\mathbf{I}, \hat{\mathbf{I}}) = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (I^G_{x,y} - \hat{I}^G_{x,y})^2 \qquad [6.29]$$

The peak signal-noise ratio on the green plane is then computed by using MSE^G in equation [6.28]. These marginal criteria are useful for assessing the degradation due to demosaicing on each component, especially on the green plane (often estimated first, thus crucial to estimate the red and blue planes).

6.4.2. *Fidelity results and discussion*

To give an insight into the performance of demosaicing methods, we select the main ten among those presented in sections 6.2 and 6.3. These are listed in the legend of Figure 6.9 and repeated in Table 6.1[9]. The images used for these tests are the commonly encountered ones in the

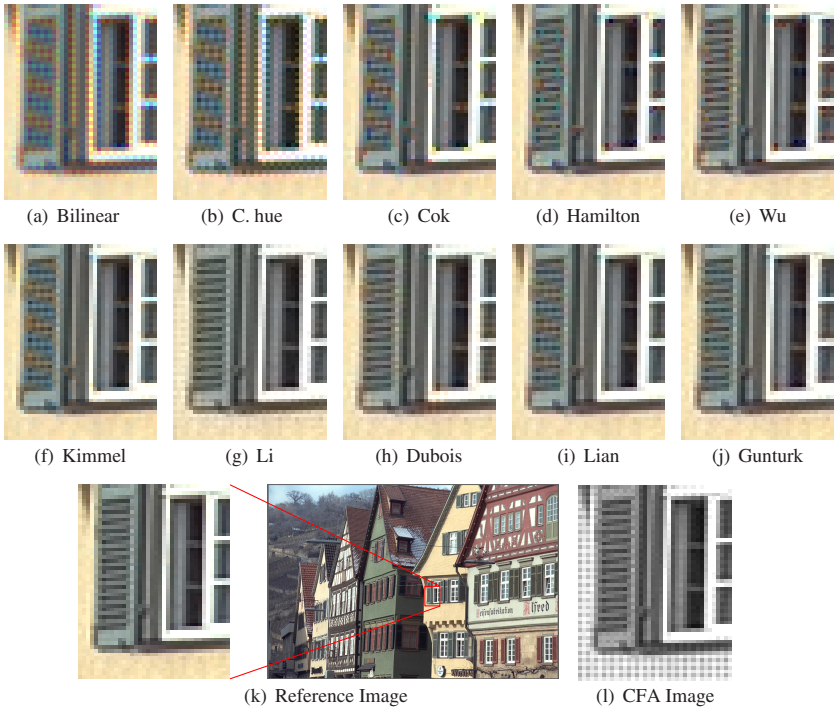
9 The source code is available online for some of these methods:

– Bilinear, C. hue, Cok and Hamilton (by Ting Chen): http://scien.stanford.edu/pages/labsite/1999/psych221/projects/99/tingchen/main.htm,

– Li, Gunturk: http://www.csee.wvu.edu/ xinl/demo/demosaic.html,

– Dubois (with [VAN 07]): http://lcavwww.epfl.ch/reproducible_research/VandewalleKAS07.

demosaicing literature, namely the twelve natural images from Kodak database [KOD 91][10].



(a) Bilinear (b) C. hue (c) Cok (d) Hamilton (e) Wu

(f) Kimmel (g) Li (h) Dubois (i) Lian (j) Gunturk

(k) Reference Image (l) CFA Image

Figure 6.9. *Demosaicing results obtained by ten demosaicing methods on a sample of the image "Houses"(k)(number 8 of Kodak database): (a) by bilinear interpolation [COK 86], (b) under the assumption of constant hue [COK 87], (c) by pattern recognition [COK 94], (d) using a gradient [HAM 97], (e) using component consistency of interpolation directions [WU 04], (f) with adaptive weighting [KIM 99], (g) using the local covariance [LI 01], (h) by frequency selection [DUB 05], (i) by frequency and spatial analyses [LIA 07], (j) by alternating projection of components [GUN 02]*

Figure 6.9 gives a visual overview of the results provided by the ten methods on a sample of the image "Houses". We can first note that all the estimated images display demosaicing artifacts. No method is

10 Kodak database is available at http://www.math.purdue.edu/ lucier/PHOTO_CD, and the selected twelve images show a significant diversity of colors and of textured regions.

able to perfectly reproduce the reference image shown in Figure 6.9k. Indeed, the presence of fine details makes this image particularly difficult to estimate: in regions of high spatial frequencies, the CFA sampling causes a spectrum aliasing, which means an inevitable loss of information under the Nyquist-Shannon theorem. Demosaicing artifacts are of three main types.

– the blurring effect is characterized by an attenuation of transitions and a loss of very fine details. This effect is especially marked for the image estimated by bilinear interpolation because this method performs a low-pass filtering (for instance, see the texture of the wall in Figure 6.9a), but it is also found in most of the other methods (see the attenuation of the transoms of the opened window side);

– the *zipper effect* is a typical demosaicing artifact. It appears as a repeating (high and low levels alternation) and colored (cyan or orange) pattern that mainly occurs along the horizontal, vertical, or diagonal transitions. The main cause of this artifact is an interpolation of levels that belong to homogeneous regions representing different objects. It is particularly noticeable in the images estimated by bilinear interpolation (Figure 6.9a) and under the assumption of constant hue (Figure 6.9b), along the uprights and the lower edge of the window;

– the *false color* artifact corresponds to a high deviation between the estimated and reference colors of a pixel. Occurring as a visually aberrant color, this effect typically derives from a loss of correlation among the color components (see section 6.2.3 and Figure 6.6b), leading in particular to a poor estimation of the components R and/or B. This artifact can be interpreted in the frequency domain as related to the mutual overlapping of the estimated luminance and chrominance spectra, due to a bad design of the selective filters used to estimate these components. All demosaicing methods tend to generate false colors, sometimes jointly with the zipper effect. Such aberrant colors are visible in the shutter slats, or at the upper-left corner of the window (notably in Figures 6.9 c, d, f and i).

The fidelity criteria described in section 6.4.1 allow us to objectively assess the demosaicing quality and to compare the relative performance of the methods. Table 6.1 shows the numerical results of fidelity obtained by the ten demosaicing methods under the *PSNR* criterion. The results

Method \ Image	6	7	8	9	11	16	19	20	21	22	23	24	Rank sum	$\overline{PSNR}$ (dB)	Operations
Bilinear	10	10	10	10	10	10	10	10	10	10	9	10	119	30.89	6
C. hue	9	9	9	9	9	9	9	9	9	9	8	8	106	33.26	11
Cok	8	7	7	7	7	8	8	7	7	7	7	6	86	34.58	43
Hamilton	5	4	5	5	4	5	5	4	5	4	5	5	56	38.20	21
Wu	4	5	4	4	5	4	4	5	4	5	4	4	52	38.62	80
Kimmel	6	6	6	6	6	6	6	6	6	6	10	6	76	35.61	300
Li	7	8	8	8	8	7	7	8	8	8	7	9	93	34.39	>1300
Dubois	1	2	1	2	1	1	2	3	2	2	2	1	20	41.16	145
Lian	3	1	3	1	3	3	3	1	3	1	1	2	25	40.49	65
Gunturk	2	3	2	3	2	2	1	2	1	3	3	3	27	40.82	384

Table 6.1. *Evaluation of the demosaicing fidelity under the peak signal-noise ratio (PSNR) criterion for twelve color images from Kodak database. In addition to the average values in the penultimate column, the table shows the performance ranks achieved by the same ten demosaicing methods as those in Figure 6.9. The last column shows the average number of operations required to estimate the color of a pixel*

under the *MSE* (whose *PSNR* is the logarithmic form) and *MAE* criteria are not reported here, but it should be noted that the latter provides very similar rankings. The relative performances of the methods are given as their rank for each image, rather than the actual *PSNR* values that are less synthetic. The average actual values are, however, given in the penultimate column of the table, and it is possible to notice that their classification coincides with that provided by the rank sum computed in the previous column.

We find that the two methods taking advantage of the frequency domain (*i.e.* those of Dubois and Lian *et al.*) consistently provide better results than methods using only a neighborhood in the image plane. Moreover, the method proposed by Dubois on the basis of the work of Alleysson *et al.* is often classified in the first two, and achieves the highest scores on average over the twelve images. Gunturk *et al.*'s method by alternate projection of components also provides excellent results. In contrast, the "historical" methods (by bilinear interpolation and using hue constancy) achieve scores well behind, which confirms the visual assessment of Figure 6.9. Moreover, globally, those rankings obtained for the twelve images coincide both with each other and roughly with the subjective quality found on image number 8 alone. A reservation

concerns the ranking of Kimmel's method for image number 23, whose highly saturated colors cause the generation of artifacts during the iterative correction phase based on a component ratio.

In the perspective of an implantation of one of these algorithms, *a fortiori* for an embedded processing of real-time demosaicing, we think it useful to add an overview of their respective complexity in the last column of Table 6.1. Due to the variety of implementations that we have (Matlab, C, etc.) and being unable to assess their optimization level, we estimate the computational complexity of the methods as the number of operations required rather than as their execution time[11]. Observed in conjunction with those of fidelity, these results suggest a quality/speed compromise. We can notably point out that Hamilton and Adams' method yields satisfactory results despite its simplicity, while Li and Orchard's method is prohibitive due to the extremely high computational cost of the matrix inversion used. Iterations, such as those integrated to Kimmel's and Gunturk *et al.*'s methods[12], are highly detrimental in terms of computational complexity. Finally, Lian *et al.*'s method probably provides the best balance between the result fidelity and the computation time.

6.5. Color camera calibration

The signals emitted by an acquisition system should be transformed appropriately to match a faithful representation of the colors as they appear to the human observer. Such a transformation is often called *calibration* and is based on a more or less sophisticated color model whose role is to make the connection between the output of the sensor(s) (after demosaicing in the case of single-sensor color cameras) and the color information of the acquired scene.

In general, such a color model is different from an acquisition system to another, mainly due to the variability of spectral sensitivities of

11 For simplicity, all the basic operations (addition, multiplication, comparison, and absolute value) are accounted equally.

12 Let us note, however, that an implementation of Gunturk *et al.*'s method requiring only one pass was recently proposed [LU 10].

the device channels and of all the nonlinearities that may arise. In addition, we should not forget that the spectral characteristics of the illuminant used is directly involved in the acquisition process. This suggests that, ideally, a different color model per illuminant should be used. Except under very controlled illumination, the huge diversity of concrete situations makes it unrealistic to have as many color models as possible acquisition configurations. In practice, most systems address the problem by adjusting the gain values of each channel during a process commonly known as *white balance*.

As suggested by its name, a white balance adjusts the output signals of the different channels depending on the spectral characteristics of the light sources, so that achromatic regions of the acquired scene are actually represented as such. For instance, with an incandescent lamp, it is necessary to increase the sensitivity of the channels dedicated to low wavelengths of the visible spectrum relative to that of channels dedicated to high wavelengths, to prevent white objects from appearing artificially yellowish.

The major difficulty in white balancing is to determine which parts of the acquired scene can be considered achromatic. The reliable way is of course to use a Lambertian test pattern of the same reflection coefficient for all wavelengths in the visible spectrum. However, if this approach is ideally suited in a professional context, it is much more difficult to implement in a general public context. In the latter case, methods are deployed to estimate the color temperature of the illuminant.

One of these methods uses a set of photodiodes typically coated with RGB filters. The photodiodes are placed in the acquisition device so that they directly receive the incident light [HAN 83]. Unfortunately, this approach may provide incorrect results due to spurious reflections from colored areas of the scene. Furthermore, additional components may induce a significant additional cost.

Better results are usually obtained by estimating the white balance adjustment from the raw data of the sensor(s). The acquired image is divided into a more or less important number of regions (typically at least 20). For each region, the average signal value of each channel is

computed. The obtained average values are then differentiated according to a given algorithm to search for specific signatures allowing to select, within a set of previously identified light sources, the most likely to illuminate the acquired scene [MOR 90, LIU 95, SPA 98]. Once the most likely source is selected, the different channels are weighted accordingly to adjust their relative contributions to the spectral distribution of energy of the incident light.

Whatever the method used to perform white balance, this operation is not in itself a complete color model [LAM 05]. The color information captured should be correlated with what an observer perceives. In other words, care should be taken to minimize the differences between digital colors and perceived colors. This refers to the problems of color appearance that has been dealt with in Chapter 3 of this book. In practical terms, the color model can be built following a white balance by performing an acquisition of a set of reference colors under the same shooting conditions as the scene to be imaged. The issue is then to determine the transformation(s) required to correlate the output values of the acquisition system with the reference colors.

The general principle is to match the observations of the selected samples with their values in a standardized color space such as *XYZ*. The underlying assumption is that the knowledge of the correspondence between a number of well-chosen colors in two spaces linked by a transformation allows us to deduce an approximation of the transformation for a richer and more extensive colors set. This implies that the so-called *Luther condition* is met, according to which a linear transformation exists between the acquisition space and the selected normalized space [HOR 84]. In practice, the Luther condition is generally not strictly fulfilled, but the general approach, however, gives satisfactory results.

The color model aims to link the values provided by the sensor(s) and the colors perceived by the observer. The values it provides can be encoded in multiple ways in color spaces that are dependent or independent of the material, depending on the applications and devices to which the acquisition system is led to be related.

6.6. Conclusion

This chapter focuses on one of the fundamental processings performed in a single-sensor color camera: demosaicing. We have seen why and how the color image is estimated from the raw data sampled from the sensor. The main principles underlying most demosaicing methods are presented here along with the outstanding algorithms that implement them. Among these, we mainly identify two families: methods based on a spatial analysis of the image, and those using the frequency domain. Other recently proposed approaches are also mentioned briefly.

The multiplicity and diversity of demosaicing methods reflect the acuteness of this problem, while the latest ones show its topicality. The steady expansion of the research on this topic for over 30 years has allowed significant progress in the quality of the estimated image, and more precisely in its fidelity to the observed scene. The brief fidelity study presented here, which is performed on a representative sample of the methods mentioned and on a representative set of images, highlights the superiority of approaches analyzing the CFA signal in the frequency domain. Indeed, these provide the most accurate results under the signal-noise ratio criterion, for a relatively low computational cost.

In the context of color image formation within single-sensor color devices, we strictly confined ourselves here to a fidelity study of the estimated image. It would be interesting to extend the quality study of this image in the context of its display or reproduction. This requires a complete control over the acquisition conditions and the use of other quality metrics, better correlated with the human visual perception, as well as subjective quality tests.

6.7. Bibliography

[ALL 05] ALLEYSSON D., SÜSSTRUNK S., HÉRAULT J., "Linear demosaicing inspired by the human visual system", *IEEE Transactions on Image Processing*, vol. 14, no. 4, p. 439-449, April 2005.

[BAY 76] BAYER B.E., "Color imaging array", U.S. patent 3,971,065, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., July 1976.

[CHA 06] CHANG L., TAN Y.-P., "Hybrid color filter array demosaicking for effective artifact suppression", *Journal of Electronic Imaging*, vol. 15, no. 1, p. 013003,1-17, January 2006.

[CHA 07] CHAIX DE LAVARÈNE B., ALLEYSSON D., HÉRAULT J., "Practical implementation of LMMSE demosaicing using luminance and chrominance spaces", *Computer Vision and Image Understanding*, vol. 107, no. 1-2, p. 3-13, July 2007.

[CHU 06] CHUNG K.-H., CHAN Y.-H., "Color demosaicing using variance of color differences", *IEEE Transactions on Image Processing*, vol. 15, no. 10, p. 2944-2955, October 2006.

[COK 86] COK D.R., "Signal processing method and apparatus for sampled image signals", U.S. patent 4,630,307, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., December 1986.

[COK 87] COK D.R., "Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal", U.S. patent 4,642,678, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., February 1987.

[COK 94] COK D.R., "Reconstruction of CCD images using template matching", *Proceedings of the IS&T's 47th Annual Conference, Physics and Chemistry of Imaging Systems (ICPS '94)*, vol. 2, p. 380-385, Rochester, New York, USA, May 1994.

[CON 10] CONDAT L., "A simple, fast and efficient approach to denoisaicking: joint demosaicking and denoising", *Proceedings of the IEEE International Conference on Image Processing (ICIP '10)*, p. 905-908, Hong Kong, China, September 2010.

[DUB 05] DUBOIS E., "Frequency-domain methods for demosaicking of Bayer-sampled color images", *IEEE Signal Processing Letters*, vol. 12, no. 12, p. 847-850, December 2005.

[FRE 88] FREEMAN W.T., "Median filter for reconstructing missing color samples", U.S. patent 4,724,395, to Polaroid Co., Patent and Trademark Office, Washington D.C., December 1988.

[GUN 02] GUNTURK B.K., ALTUNBASAK Y., MERSEREAU R.M., "Color plane interpolation using alternating projections", *IEEE Transactions on Image Processing*, vol. 11, no. 9, p. 997-1013, September 2002.

[HAM 97] HAMILTON J.F., ADAMS J.E., "Adaptive color plan interpolation in single sensor color electronic camera", U.S. patent 5,629,734, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., May 1997.

[HAN 83] HANMA K., MASUDA M., NABEYAMA H., SAITO Y., "Novel technologies for automatic focusing and white balancing of solid state color video camera", *IEEE Transactions on Consumer Electronics*, vol. CE-29, no. 3, p. 376-382, August 1983.

[HIB 95] HIBBARD R.H., "Apparatus and method for adaptively interpolating a full color image utilizing luminance gradients", U.S. patent 5,382,976, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., January 1995.

[HIR 05] HIRAKAWA K., PARKS T.W., "Adaptive homogeneity-directed demosaicing algorithm", *IEEE Transactions on Image Processing*, vol. 14, no. 3, p. 360-369, March 2005.

[HOR 84] HORN B.K.P., "Exact reproduction of colored images", *Computer Vision, Graphics, and Image Processing*, vol. 26, no. 2, p. 135-167, May 1984.

[KAP 00] KAPAH O., HEL-OR Z., "Demosaicing using artificial neural networks", *Proceedings of the SPIE - Applications of Artificial Neural Networks in Image Processing V*, vol. 3962, p. 112-120, San Jose, CA, USA, January 2000.

[KER 99] KEREN D., OSADCHY M., "Restoring subsampled color images", *Machine Vision and Applications*, vol. 11, no. 4, p. 197-202, December 1999.

[KIM 99] KIMMEL R., "Demosaicing: image reconstruction from color CCD samples", *IEEE Transactions on Image Processing*, vol. 8, no. 9, p. 1221-1228, September 1999.

[KOD 91] KODAK ET DIVERS PHOTOGRAPHES, "Kodak photo CD PCD0992, access software & photo sampler, final version 2.0", [CD-ROM, Part No. 15-1132-01], 1991.

[LAM 05] LAM E.Y., "Combining gray world and retinex theory for automatic white balance in digital photography", *Proceedings of the 9th IEEE International Symposium on Consumer Electronics (ISCE '05)*, Hong Kong, China, p. 134-139, June 2005.

[LAR 93] LAROCHE C.A., PRESCOTT M.A., "Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients", U.S. patent 5,373,322, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., June 1993.

[LI 00] LI X., Edge directed statistical inference and its applications to image processing, PhD thesis, Princeton University, New Jersey, USA, November 2000.

[LI 01] LI X., ORCHARD M.T., "New edge-directed interpolation", *IEEE Transactions on Image Processing*, vol. 10, no. 10, p. 1521-1527, October 2001.

[LI 08] LI X., GUNTURK B.K., ZHANG L. "Image demosaicing: a systematic survey", PEARLMAN W.A., WOODS J.W., LU L. (eds), *Proceedings of the SPIE Conference on Visual Communications and Image Processing (VCIP '08)*, vol. 6822, San Jose, California, USA, January 2008.

[LIA 05] LIAN N., CHANG L., TAN Y.-P., "Improved color filter array demosaicking by accurate luminance estimation", *Proceedings of the 12th International Conference on Image Processing (ICIP '05)*, vol. 1, Geneva, Italy, p. I-41-4, September 2005.

[LIA 07] LIAN N.-X., CHANG L., TAN Y.-P., ZAGORODNOV V., "Adaptive filtering for color filter array demosaicking", *IEEE Transactions on Image Processing*, vol. 16, no. 10, p. 2515-2525, October 2007.

[LIU 95] LIU Y.-C., CHAN W.-H., CHEN Y.-Q., "Automatic white balance for digital still camera", *IEEE Transactions on Consumer Electronics*, vol. 41, no. 3, p. 460-466, August 1995.

[LON 02] LONGERE P., XUEMEI Z., DELAHUNT P.B., BRAINAR D.H., "Perceptual assessment of demosaicing algorithm performance", *Proceedings of the IEEE*, vol. 90, no. 1, p. 123-132, January 2002.

[LOS 10] LOSSON O., MACAIRE L., YANG Y., "Comparison of color demosaicing methods", *Advances in Imaging and Electron Physics*, vol. 162, p. 173-265, July 2010.

[LU 03] LU W., TAN Y.-P., "Color filter array demosaicking: new method and performance measures", *IEEE Transactions on Image Processing*, vol. 12, no. 10, p. 1194-1210, October 2003.

[LU 10] LU Y., KARZAND M., VETTERLI M., "Demosaicking by alternating projections: theory and fast one-step implementation", *IEEE Transactions on Image Processing*, vol. 19, no. 8, p. 2085-2098, August 2010.

[LUK 05] LUKAC R., PLATANIOTIS K.N., "Universal demosaicking for imaging pipelines with an RGB color filter array", *Pattern Recognition*, vol. 38, p. 2208-2212, April 2005.

[LUK 08] LUKAC R., *Single-Sensor Imaging: Methods and Applications for Digital Cameras*, Image Processing Series, CRC Press, September 2008.

[MEN 06] MENON D., ANDRIANI S., CALVAGNO G., "A novel technique for reducing demosaicing artifacts", *Proceedings of the XIVth European Signal Processing Conference (EUSIPCO '06)*, Florence, Italy, September 2006.

[MEN 09] MENON D., CALVAGNO G., "Regularization approaches to demosaicking", *IEEE Transactions on Image Processing*, vol. 18, no. 10, p. 2209-2220, October 2009.

[MOR 90] MORIMURA A., UOMORI K., KITAMURA Y., FUJIOKA A., HARADA J., IWAMURA S., HIROTA M., "A digital video camera system", *IEEE Transactions on Consumer Electronics*, vol. 36, no. 4, p. 3866-3876, November 1990.

[MUK 01] MUKHERJEE J., PARTHASARATHI R., GOYAL S., "Markov random field processing for color demosaicing", *Pattern Recognition Letters*, vol. 22, no. 3-4, p. 339-351, March 2001.

[RIB 08] RIBÉS A., SCHMITT F., "Linear inverse problems in imaging", *IEEE Signal Processing Magazine*, vol. 25, no. 4, p. 84-99, July 2008.

[SOU 09] SOULEZ F., THIÉBAUT E., "Joint deconvolution and demosaicing", *Proceedings of the IEEE International Conference on Image Processing (ICIP '09)*, Le Caire, Egypt, p. 145-148, November 2009.

[SPA 98] SPAULDING K.E., VOGEL R.M., SZCZEPANSKI J.R., "Method and apparatus for color-correcting multi-channel signals of a digital camera", US patent 5,805,213, to Eastman Kodak Co., Patent and Trademark Office, Washington D.C., September 1998.

[TAU 00] TAUBMAN D.S., "Generalized Wiener reconstruction of images from colour sensor data using a scale invariant prior", *Proceedings of the 7th International Conference on Image Processing (ICIP 2000)*, vol. 3, Vancouver, BC, Canada, p. 801-804, September 2000.

[VAN 07] VANDEWALLE P., KRICHANE K., ALLEYSSON D., SÜSSTRUNK S., "Joint demosaicing and super-resolution imaging from a set of unregistered aliased images", *Proceedings of the 19th IST/SPIE Electronic Imaging Annual Symposium (SPIE '07)*, vol. 6502 of *Digital Photography III*, p. 65020A.1-65020A.12, San Jose, California, USA, January 2007.

[WU 04] WU X., ZHANG N., "Primary-consistent soft-decision color demosaicking for digital cameras", *IEEE Transactions on Image Processing*, vol. 13, no. 9, p. 1263-1274, September 2004.

[YAN 09] YANG Y., Contribution à l'évaluation objective de la qualité d'images couleur estimées par dématriçage, PhD thesis, University of Lille 1, Sciences et Technologies, October 2009.

[ZHA 97] ZHANG X., WANDELL B.A., "A spatial extension of CIELAB for digital color reproduction", *Journal of the Society for Information Display*, vol. 5, no. 1, p. 61-63, March 1997.